Data Mining and Concept Clustering in Determining the Nature of a Network Attack

Chet Maciag
Air Force Research Laboratory, AFRL/IFGB
525 Brooks Rd., Rome, NY 13441-4505, United States

Information Assurance for Information Warfare [IA/IW] is defined as, "information operations that protect and defend information and information systems by ensuring their availability, integrity, authentication, confidentiality, and non-repudiation. Information assurance includes providing for restoration of information systems by incorporating protection, detection, and reaction capabilities. (DODD S-3600.1)

Information Assurance is necessary to provide commanders with the capability to defend information flows required to execute assigned missions in both peacetime and crisis/contingency. This is a 365-day-a year Information Assurance for daily operations and business at all levels, and to integrate it's provision into Operations planning and execution.

The functional state-of-the-art in Information Assurance is analogous to air-traffic control—operators continually scanning networks for signs of attack. This obviously needs improvement. reports of successful and feared-successful attacks appear with increasing frequency.

"Russian Hackers Steal US Weapons Secrets"

"American officials believe Russia may have stolen some of the nation's most sensitive military secrets, including weapons guidance systems and naval intelligence codes, in a concerted espionage offensive that investigators have called operation Moonlight Maze.

This was so sophisticated and well co-ordinated that security experts trying to build ramparts against further incursions believe America may be losing the world's first 'cyber war'.

(Interview with Mr. John Hamre,
Deputy Secretary of Defense)
London Sunday Times, 25 July 1999

TTCP TP-11 have mounted an IA project. The first year demonstrated Successful exchange of intrusion event data between Australian Shape-Vector and Rome Labs [AFRL]'s EPIC^2 prototypes

The USAF Enterprise Defence project is intended to develop the next-generation enterprise defence framework for AF Modicums and aerospace expeditionary forces (AEF), providing situational assessment and decision support, simultaneously improving the information overload problem for network defenders, provide a consistent visual environment for information portrayal, fusing information assurance (IA) and network management data into a common enterprise picture (CEP). A further intent is to empower the MAJCOM to validate and influence present and future technology so it suitable for transition into NMS/BIP and other acquisition programs

The Air Force has embarked on a project to develop the next-generation enterprise defence framework for AF MAJCOMs and Aerospace Expeditionary Forces (AEF), including situational assessment & decision support;

1. Reduce network defenders' information overload;
2. Provide a consistent visual environment for information portrayal ;
3. Fuse information assurance (IA) and network management data into a common enterprise picture (CEP); and to
4. Empower the MAJCOM to validate and influence present and future technology so it suitable for transition into NMS/BIP and other acquisition programs.

Several programs are under way to address these priorities. The presentation slides show several of these in considerable detail. There remain important potential problems for data fusion engines to solve:

- Identifying low, slow mapping and probing attempts. Sensor data grows quickly and it is difficult to store, problems with storage and retrieval; the current plan is to utilize a trend database that saves suspicious events and compressing other data.
- Acquiring knowledge from domain experts for data analysis. Some data gathering has been done but the information has not been readily available.
- Data correlation between sensors and events in real-time is needed in order to identify attacks and reduce false alarms. Throughput (for real time operation) is biggest problem. The current plan is to Implement "rules" in native code
- Goal-seeking to determine the purpose of an attack. [This will require a flexible, backward chaining capability.]
- Better rule/filter deconfliction are needed between components. That is, there is a need to ensure that filtering/rules do not conflict with each other and that a filter does not block data needed by a rule.
- Better data mining tools and techniques are needed to identify new attack signatures
- Modification of KB knowledge space must be made possible by non-KB experts, or their information and experience will be lost.
- Threat profile/identification extrapolation—this is needed to face future, potential threats and attacks.
- Machine learning algorithms are needed to enable the system to anticipate analysts "next move"

Technology assessment can examine new applications' functional goals and structures, identify the cognitively demanding aspects of decision makers' tasks, analyze work domain constraints and task context, support team decision making and co-ordination, and support software design.

Paper presented at the RTO IST Workshop on "Multimedia Visualization of Massive Military Datasets", held in Quebec, Canada, 6-9 June 2000, and published in RTO-MP-050.

Discussion – Paper 18

Data Mining and Concept Clustering in Determining the Nature of a Network Attack

Commonly most effort within the Information Assurance (IA) arena has been focused at finding the attack signature but protection, detection and reaction all need to be visualised.

As humans and machines are tuned in to see specific patterns there is a need to visualise data/information in different ways to see odd or unusual patterns. The results of one visualisation should then be able to be added into another visualisation giving the user another chance to gain more understanding from the data.

Within network management very little work has been done on the correlation between the cyber and the real world in an operational environment. It is important to do this in order to show the way that the availability of the communications network (both blue and red) could affect the operation.

In order to do this the operations communications structure has to be mapped. This includes both the hardware and how networks are connected to each other e.g. via other countries networks (whether civil or military) for both the blue and red forces. It also needs to highlight the capacity and real structure of the red theatre communication links and how they can be effected.

The aim is to fuse IA and network management into a common enterprise picture.